

Structural basis for selective inhibition of Src family kinases by PP1

Yi Liu¹, Anthony Bishop¹, Laurie Witucki¹, Brian Kraybill¹, Eiji Shimizu², Joe Tsien², Jeff Ubersax³, Justin Blethrow³, David O Morgan³ and Kevan M Shokat^{1,2*}

Background: Small-molecule inhibitors that can target individual kinases are powerful tools for use in signal transduction research. It is difficult to find such compounds because of the enormous number of protein kinases and the highly conserved nature of their catalytic domains. Recently, a novel, potent, Src family selective tyrosine kinase inhibitor was reported (PP1). Here, we study the structural basis for this inhibitor's specificity for Src family kinases.

Results: A single residue corresponding to Ile338 (v-Src numbering; Thr338 in c-Src) in Src family tyrosine kinases largely controls PP1's ability to inhibit protein kinases. Mutation of Ile338 to a larger residue such as methionine or phenylalanine in v-Src makes this inhibitor less potent. Conversely, mutation of Ile338 to alanine or glycine increases PP1's potency. PP1 can inhibit Ser/Thr kinases if the residue corresponding to Ile338 in v-Src is mutated to glycine. We have accurately predicted several non-Src family kinases that are moderately ($IC_{50} \sim 1 \mu M$) inhibited by PP1, including c-Abl and the MAP kinase p38.

Conclusions: Our mutagenesis studies of the ATP-binding site in both tyrosine kinases and Ser/Thr kinases explain why PP1 is a specific inhibitor of Src family tyrosine kinases. Determination of the structural basis of inhibitor specificity will aid in the design of more potent and more selective protein kinase inhibitors. The ability to desensitize a particular kinase to PP1 inhibition of residue 338 or conversely to sensitize a kinase to PP1 inhibition by mutation should provide a useful basis for chemical genetic studies of kinase signal transduction.

Introduction

Protein kinases catalyze the transfer of the γ -phosphate from ATP to serine, threonine, tyrosine or histidine residues of protein substrates. They play a central role in controlling many signal transduction pathways including those in B and T cells of the immune system and neurons in the brain [1–4]. In fact, all eukaryotic cells contain many protein kinase regulatory systems (it is estimated that 2% of the human genome encodes protein kinases [5]). Because of the enormous size of the protein kinase family, the tremendous redundancy and overlapping substrate specificities of many protein kinases, the dissection of kinase signaling pathways is a challenge to existing genetic and biochemical techniques [6]. For example, kinase gene-knockouts often show no phenotype because other cellular kinases compensate for the missing kinase activity [7–9]. One approach that avoids the problems associated with gene compensation is the development of mono-specific small-molecule inhibitors of engineered kinases [10,11]. Such inhibitors could be used to selectively block kinase function in order to map pathways. Kinase inhibitors are

also potential therapeutics, but tremendous effort is required to identify compounds that can selectively bind to one particular kinase out of the estimated 2000 highly homologous human protein kinases [12–14].

One simplifying approach is to develop 'family-selective' protein kinase inhibitors that can potently inhibit all highly homologous members of a protein kinase family. The sequence identity of the kinase catalytic domains is high for kinases within a family (70–85% within Src family members), but is smaller between families (~20% between c-Src and protein kinase A). One protein kinase family for which specific inhibitors have been reported is the Src family of protein kinases. The Src family of tyrosine kinases contains nine members, Src, Lck, Fyn, Lyn, Hck, Yes, Fgr, Blk and Yrk, which are involved in lymphocyte development (Lck and Fyn) [2,15], platelet activation (c-Src, Fyn and Lyn) [16,17], mast cell degranulation (Lyn, c-Src and Yes) [18,19] and many other cell functions. One reason for the intense interest in inhibitors of Src family kinases is their involvement in regulating immune cell functions. Antigen stimulation of T cells

Addresses: ¹Department of Chemistry and ²Department of Molecular Biology, Princeton University, Princeton, NJ 08544, USA. ³Department of Physiology, University of California, San Francisco, CA 94143-0444, USA.

*Present Address: Department of Cellular and Molecular Pharmacology, Box 0450, UC San Francisco, San Francisco, CA 94143-0450, USA.

Correspondence: Kevan M Shokat
E-mail: shokat@cmp.ucsf.edu

Key words: kinase inhibitor, modeling, PP1, selective, Src

Received: 24 April 1999
Revisions requested: 18 May 1999
Revisions received: 24 June 1999
Accepted: 25 June 1999

Published: 13 August 1999

Chemistry & Biology September 1999, 6:671–678
<http://biomednet.com/elecref/1074552100600671>

1074-5521/99/\$ – see front matter
© 1999 Elsevier Science Ltd. All rights reserved.

results in increased enzymatic activity of Fyn [20]. Conversely, overexpression of a catalytically inactive form of Fyn (Fyn (Lys296→Glu)) in thymocytes of transgenic mice substantially inhibited TCR-mediated T cell activation [21]. The latter genetic evidence for the essential role of Fyn in T cell activation has made it a particularly attractive drug target for development of immunosuppressants with new modes of action.

Several Src family selective inhibitors have been described [22–24], including 4-amino-1-*tert*-butyl-3-(*p*-methylphenyl) pyrazolo[3,4-*d*]pyrimidine (PP1) [24]. PP1 has gained the most widespread use because of its potency and ease of synthesis [25–28]. Hanke *et al.* [24] showed that PP1 inhibits Lck and Fyn *in vitro* at concentrations significantly lower than those required to inhibit non-Src family kinases ZAP-70, JAK2 and protein kinase A present in T cells. PP1 is also active in cells as demonstrated by the inhibition of whole cell tyrosine phosphorylation and proliferation in T cells stimulated with anti-CD3 antibody and mitogens, as well as B cells [29]. PP1 can therefore be used to dissect the role of the Src family component of T cell receptor signaling and provides a complement to other T cell signaling blockers such as FK506, cyclosporin, rapamycin and wortmanin [30–32].

Two key problems remain with the use of PP1 as a tool for deconvoluting Src family dependent cell signaling pathways. First, PP1 shows little specificity between Src family kinases and therefore cannot be used to dissect the role of individual Src family kinases that are coexpressed in a given cell (e.g. T cells, B cells, mast cells and dendritic cells). Second, the assertion that PP1 is truly Src family selective has not been rigorously tested.

The problem of rigorously checking whether a given kinase inhibitor can inhibit other ('off-target') kinases in the organism of interest is common to all kinase inhibitors. The reason for this is that not all human protein kinases have been identified. Furthermore, the task of expressing, purifying and assaying anything close to all human kinases (~2000) is an almost impossible task. New sequence homology based methods coupled to high-throughput screening for kinase inhibitors might obviate the need to screen against all naturally occurring protein kinases [33]. Yet without such structure–activity data the use of small-molecule inhibitors in cell-signaling studies may be misleading. To address this problem we have identified the key binding site amino acid residue responsible for imparting selective kinase inhibition of PP1. Through systematic mutagenesis of Ile338 in *v*-Src, as well as the non-Src family kinases *c*-Abl, CaMKII α , Cdk2 and p38, we have found a strong correlation between PP1 binding affinity and amino acid sidechain volume at the position corresponding to Ile338 in *v*-Src. We used this sequence information to identify

another non-Src family kinase (*c*-Abl) that is potently inhibited by PP1. This study shows that a large component of PP1 selectivity results from residue 338 and that other protein kinases can be analyzed in the same way to reveal the potential for inhibition by PP1.

Results and discussion

Molecular modeling of PP1 binding to Hck and comparison with kinase sequence alignments

As reported by Hanke *et al.* [24] PP1 is a potent inhibitor of Lck ($IC_{50} = 0.005 \mu M$) and Fyn ($IC_{50} = 0.006 \mu M$), but is much less potent for ZAP-70 ($IC_{50} > 100 \mu M$) and JAK2 ($IC_{50} > 50 \mu M$; Table 1). To date, the crystal structure of PP1 bound to a kinase has not been solved. We therefore used the co-crystal structure of the Src family kinase Hck bound to the general kinase inhibitor quercetin and a non-hydrolyzable ATP analog AMP–PNP as a guide to dock PP1 into the ATP-binding site of Hck (Figure 1a–d) [34]. From this predicted binding mode, the C3 phenyl ring is in close contact with the sidechain of residue Thr338. This proximity of the C3 phenyl ring and the Thr338 sidechain suggested to us that this residue could be a key determinant for PP1 binding to Src family kinases.

The position adjacent to C3, the N4 exocyclic amine of PP1, has already been shown to be in close contact with Ile338. This binding mode was confirmed through synthesis of PP1 analogs with substituents at the N4 position, and design of compensatory mutations (Ile338→Ala or Ile338→Gly) in *v*-Src [10]. When a bulky group such as (*p*-*tert*-butyl) benzoyl is added to the N4 amine of PP1, this PP1 analog does not inhibit wild-type Src family kinases ($IC_{50} > 300 \mu M$ for wild-type *v*-Src and wild-type Fyn) and it inhibits the Ile338→Gly *v*-Src mutant ($IC_{50} = 0.43 \mu M$) and Thr339→Gly Fyn mutant ($IC_{50} = 0.83 \mu M$) potently. Further molecular modeling suggested that the C3 phenyl ring adjacent to N4 amine of PP1 is also in close contact

Table 1

IC_{50} values of PP1 inhibition for Lck, Fyn, ZAP-70 and JAK2 protein tyrosine kinases.

| Tyrosine kinase | IC_{50} (μM) |
|-----------------|-----------------------|
| Lck | 0.005–0.001 |
| Fyn | 0.006–0.001 |
| ZAP-70 | > 100 |
| JAK2 | > 50 |

Data taken from Hanke *et al.* [24].

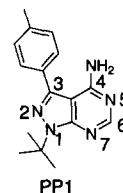
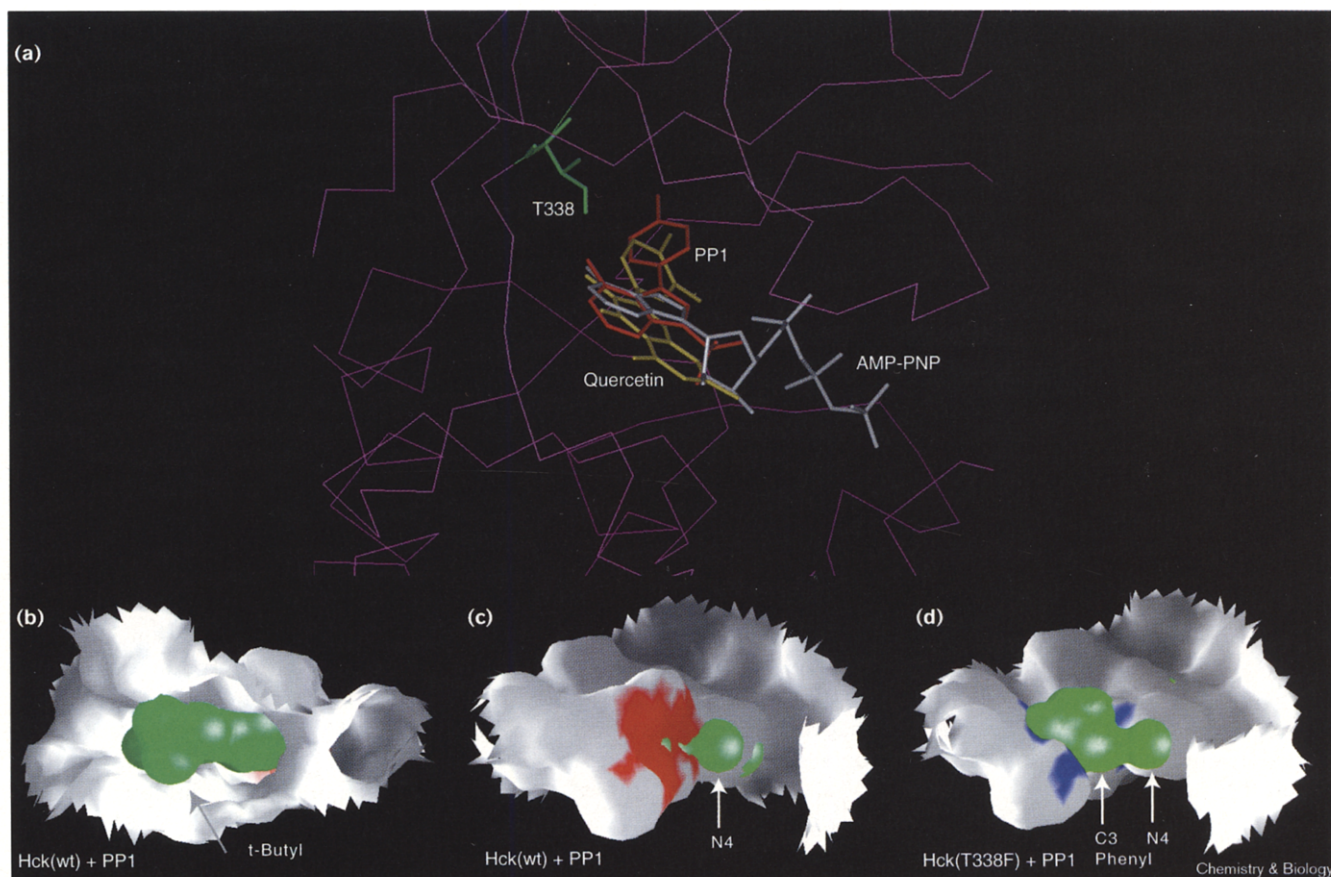


Figure 1



(a) Predicted binding orientation of PP1 (red) in the Src family tyrosine kinase Hck active site. The crystal structures of Hck bound to AMP-PNP (white) and Hck bound to quercetin (yellow) were superimposed according to the Hck protein backbone (purple). The orientation of PP1 was manually docked into the kinase active site by minimizing the van der Waals interaction between PP1 and Hck using the dock program of InsightII. **(b–d)** Proposed structural basis for the PP1 inhibition of Src family kinases as shown by a surface representation of the ATP-binding pocket of the Src family tyrosine

kinase Hck bound to PP1 and the proposed effect of mutating Thr338→Phe. **(b)** Hck (wt) + PP1 (front view). The solvent-accessible surface of Hck within 5 Å of PP1 is shown in gray; the surface of PP1 is shown in green. The binding orientation of PP1 is the same as in (a). **(c)** Hck (wt) + PP1 (back view). The portion of Hck's surface formed by Thr338 is shown in red. **(d)** Hck (Thr338→Phe) + PP1. Model of the unfavorable packing of PP1 into Hck (Thr338→Phe) binding site, showing the steric clash between the C3 phenyl ring and the sidechain of Phe338 (blue).

with the residue corresponding to Ile338 in v-Src [11]. These C3 modified PP1 derived inhibitors were shown to be competitive inhibitors with respect to ATP [11].

Most Ser/Thr kinases have larger hydrophobic residues (Met, Phe, Leu and Gln) at the position corresponding to Thr338 in the Src family kinases (Figure 2). From a simple docking analysis it appears that such large amino acid sidechains would sterically clash with the phenyl ring in PP1 (Figure 1d). We predicted that PP1 does not inhibit Ser/Thr kinases because these kinases contain a larger residue than threonine at the position corresponding to 338 in c-Src. Sequence alignments of the nine members of the Src family tyrosine kinases support this idea in that they all have a threonine at position 338 (Figure 2). In contrast, each of the noninhibited kinases

reported by Hanke *et al.* [24], ZAP-70, JAK2 and PKA, all have a methionine at this position (Figure 2). On the basis of this analysis we propose that the position corresponding to Thr338 in Hck controls the binding of PP1 to other kinases. Non-Src family kinases that have a threonine at the position corresponding to 338 in the Src family, such as c-Abl, PDGFR and p38, could also be potentially inhibited by PP1.

Effect of position 338 amino acid identity on PP1 inhibition

The first experimental evidence that PP1 derived its ability to inhibit Src family kinases because this family of kinases shared a threonine residue at position 338 (Src numbering will be used throughout) came from the discovery that PP1 was a 700 nM inhibitor of c-Src and yet was a poor ($IC_{50} = 5 \mu M$) inhibitor of the almost identical

Figure 2

| | | 338 | | IC ₅₀ , PP1 <1 μM |
|---------|-------|--|--|------------------------------|
| v-Src | (318) | RHEKLVQLYAMVSE-----EPIYIVTEYMSK--GSLLDFLKGGEMGY | | - |
| c-Src | (318) | RHEKLVQLYAVVSE-----EPIYIVTEYMSK--GSLLDFLKGETGKY | | + |
| Lck | (296) | QHQLVRLYAVVTQ-----EPIYIVTEYMN--GSLVDFLKTSPGKIK | | + |
| Fyn | (319) | KHDKLVQLYAVVSE-----EPIYIVTEYMNK--GSLLDFLKDGEGRA | | + |
| c-Yes | (325) | RHDKLVPLLYAVVSE-----EPIYIVTEFMSK--GSLLDFLKGGDGKY | | |
| Yrk | (318) | RHDKLVQLYAVVSE-----EPIYIVTEFMSQ--GSLLDFLKGGDGRY | | |
| c-Fgr | (311) | RHDKLVQLYAVVSE-----EPIYIVTEFMCH--GSLLDFLKNPEQGD | | |
| Lyn | (295) | QHDKLVRLYAVVTR-----EPIYIVTEYMAK--GSLLDFLKSDEGGK | | |
| Hck | (318) | QHDKLVKLHVVTK-----EPIYIVTEYMAK--GSLLDFLKSDEGSK | | |
| Blk | (287) | QHERLVRLYAVVTR-----EPIYIVTEYMAR--GCLLDFLKTDEGSR | | |
| Abl | (313) | KHPNLVQLLGVCTRE-----PPFYIVTEFMTY--GNLLDYLRRCNRQE | | + |
| Btk | (473) | SHEKLVQLYGVCTKQ-----RPIFIVTEYMAN--GCLLNLYLRMRHR | | |
| Csk | (244) | RHSNLVQLLGVVVEEK-----GGLYIVTEYMAK--GSLVDYLRSRVS | | |
| PDGFR | (660) | PHLNVVNLGACTKG-----GPIYIVTEYCRY--GDLVDYLRHNKHTF | | |
| p38 | (85) | GLLDVFTPARSLEEF-----NDVVLVTHLMGA--DLNNIVKQCQLTDD | | + |
| ZAP-70 | (394) | DNPIVRLIGVCA-----EALMLVEMAGG--GPLHKFL-VGKREE | | - |
| JAK2 | (906) | QHDNIVKYKVCYSAGR-----RNLRLIMEYLPY--GSLRDYLRQHKER | | - |
| PKA | (99) | NFPFLVKLEFSFKDN-----SNLYMVMYVPG--GEMFSLRRIGR | | - |
| CamK II | (68) | KHPNIVRLHDSISEE-----GHHYLIFDLVTG--GELFEDIVAREY | | - |
| Cdk2 | (59) | NHPNIVKLLDVIHTE-----NKLYLVTEFLHQ--DLKKFMDASALTG | | - |

Sequence alignment of residues surrounding Ile338 (v-Src) in Src family tyrosine kinases and non-Src family tyrosine and Ser/Thr kinases. The numbering of the starting residue of the sequence is listed in the parentheses before the first residue of each sequence. The residue corresponding to Ile338 in v-Src is highlighted in yellow. +, kinases inhibited by PP1 with IC₅₀ < 1 μM; -, kinases with IC₅₀ > 1 μM; the inhibition of the remainder of the kinases was not determined.

oncogenic protein v-Src. v-Src and c-Src are almost identical in their catalytic domains (98% identical). The only amino acid substitution in v-Src that maps to the active site is a Thr338→Ile substitution. As we found that PP1 is not a potent inhibitor for v-Src (Ile338) but is potent against c-Src (Thr338), it suggested to us that the modest Thr→Ile mutation at position 338 might be essential for potent PP1 inhibition.

To more carefully test the relationship between residue 338 and PP1 potency, we mutated Ile338 in v-Src to eight different amino acids (Phe, Met, Thr, Val, Cys, Ser, Ala and Gly). A truncated form of v-Src that contains the catalytic domain of v-Src, but no functional SH3 or SH2 domains (XD4) was used as the model kinase in this study because it can be expressed in *Escherichia coli*, whereas other Src family kinases must be expressed in Sf9 insect cells. The XD4 fragment was fused to glutathione S-transferase (GST), which has been shown to aid in expression of protein kinases in bacteria [35,36]. The dimeric state of GST itself does not interfere with protein kinase activity as the GST-fused kinase exhibits the same kinetic constants as v-Src without the GST fusion [37]. When Ile338 was mutated to larger residues (phenylalanine and methionine), these v-Src mutants, Ile338→Phe XD4 and Ile338→Met XD4, were poorly inhibited by PP1 (IC₅₀ = 8 μM for both mutants). The percent activity of Ile338→Met and Ile338→Phe XD4 mutants are about 25% and 10% of that of wild-type XD4, respectively (Table 2), and are comparable with the activity of Ile338→Thr mutant. It is surprising that position 338 can tolerate almost any sized sidechain, without complete loss of catalytic activity. On the basis of these results with v-Src, we expect that the Thr338→Phe or Thr338→Met mutants of c-Src could potentially be useful in studies where a PP1 insensitive c-Src (or other Src family kinase) is needed.

Is a small residue at position 338 sufficient to confer PP1 inhibition?

To test whether a small amino acid at position 338 is sufficient for potent PP1 inhibition, we mutated Ile338 to threonine, serine, alanine and glycine in XD4. The IC₅₀ value of Ile338→Thr XD4 (100 nM) is 50 times lower than the wild-type XD4. The smaller substitution (Ile338→Ser) afforded a more poorly inhibited kinase (IC₅₀ = 0.4 μM), perhaps because of the hydrophilic character of serine compared with threonine and the predicted preference of the C3 phenyl ring to bind in a hydrophobic pocket. The most potently inhibited v-Src mutants contained the smallest naturally occurring residues, alanine and glycine (IC₅₀ = 0.005 μM for Ile338→Ala and Ile338→Gly mutants). Mutation of Ile338 to alanine therefore results in 1000-fold more potent kinase inhibition by PP1. To test the generality of this mutation in other Src family kinases, we also mutated Thr339 (corresponding to Ile338 in v-Src) in Fyn to alanine. PP1 inhibits wild-type Fyn with an IC₅₀ of 0.05 μM, and is tenfold more potent against Thr338→Ala Fyn (IC₅₀ = 0.005 μM).

On the basis of the structure-activity relationships of PP1 with v-Src and Fyn mutants at position 338, we conclude that PP1 is Src family selective because all cellular Src family members have a threonine at position 338. Based on this analysis we asked if other kinases might be inhibited by PP1 that have not been reported. We reasoned that PP1 might effectively (IC₅₀ < 1.0 μM) inhibit non-Src family tyrosine kinases that have a threonine at the position corresponding to 338 in c-Src. The tyrosine kinase c-Abl has a threonine at position 334 (corresponding to 338 in v-Src; Figure 2). We used a truncated c-Abl containing only the intact SH2 and SH1 domains that is active when expressed in *E. coli* [38] and determined that PP1 inhibits the non-Src family kinase, c-Abl, with an IC₅₀ of 0.3 μM. Compared

with the Ile338→Thr v-Src mutant ($IC_{50} = 0.1 \mu\text{M}$) and wild-type Fyn ($IC_{50} = 0.05 \mu\text{M}$), PP1 is only threefold less potent against the non-Src family kinase c-Abl. The finding that a non-Src family kinase, c-Abl (50% identical to Src family kinases) can be inhibited by PP1, purely based on a sequence alignment comparison of one residue in the kinase binding site, confirms that residue 338 is essential for PP1 binding to protein kinases. Furthermore, PP1 may exhibit similar binding affinities for kinases with very divergent sequences from the Src family yet which contain a threonine or serine at the position corresponding to position 338 in c-Src (such as PDGFR α & β , c-kit, Csk and p38; Figure 2).

Mutation of position 338 to alanine or glycine to sensitize Ser/Thr kinases to PP1 inhibition

To test the possibility that PP1 could inhibit members of the Ser/Thr kinase family of enzymes if they contained a small residue at position 338 we mutated several highly divergent kinases to contain an alanine or glycine at this position. The Ser/Thr kinases in general have a large hydrophobic residue at the site corresponding to position 338 in c-Src. This fact suggests that PP1 will not be a very effective inhibitor of these kinases. Furthermore, there may be other amino acid differences responsible for poor inhibition by PP1. We tested this using the Ser/Thr kinase Ca^{2+} /calmodulin dependent kinaseII α (CamKII α) and cyclin-dependent kinase 2 (Cdk2). CamKII regulates diverse cellular processes, including neurotransmitter release, muscle contraction, gene expression and cell proliferation [39]. Cdk2 is an important regulator of the timing and co-ordination of eukaryotic cell-cycle events [40]. Both CamKII α and Cdk2 have a phenylalanine at the position (Phe89 in CamKII α and Phe80 in Cdk2) corresponding to Ile338 in v-Src. We expressed CamKII α as a GST fusion protein in *E. coli* and expressed Cdk2 in insect cells. PP1 shows very poor inhibition of wild-type CamKII α ($IC_{50} = 60 \mu\text{M}$) and wild-type Cdk2 ($IC_{50} = 50 \mu\text{M}$) as predicted on the basis of the requirement for a small sidechain at position 338. When Phe89 is mutated to alanine, the Phe89→Ala CamKII α possessed an $IC_{50} = 10 \mu\text{M}$ for PP1. When Phe89 is further mutated to glycine, the mutant displays an inhibition $IC_{50} = 0.5 \mu\text{M}$, only 5–10-fold lower than inhibition of wild-type Src family kinases. The Phe80→Gly Cdk2 mutant is even more potently inhibited ($IC_{50} = 0.16 \mu\text{M}$), representing a > 300-fold improvement in PP1 inhibition.

Although most Ser/Thr kinases have a larger residue than threonine at the position corresponding to Ile338 in v-Src, a mitogen-activated protein kinase (MAPK), p38, contains a threonine at this position (Thr106). Thr106 in p38 has already been shown to control the sensitivity of a MAP kinase inhibitor SB203580 [41–43]. SB203580 potently inhibits MAP kinases that have a threonine at position 106 such as p38, and poorly inhibits those with a residue larger

Table 2

IC_{50} values of PP1 inhibition for XD4, Fyn, Abl, CamKII, Cdk2 wild-type and mutant kinases.

| Kinase | IC_{50} (μM) | % Activity (relative to wild type) |
|------------------------|-----------------------------|------------------------------------|
| Tyrosine kinase | | |
| GST-XD4 (Ile338→Phe) | 8 ± 2 | 10 |
| GST-XD4 (Ile338→Met) | 8 ± 2 | 25 |
| GST-XD4 (wt) | 5 ± 1 | 100 |
| GST-XD4 (Ile338→Thr) | 0.1 ± 0.02 | 30 |
| GST-XD4 (Ile338→Val) | 0.1 ± 0.02 | 60 |
| GST-XD4 (Ile338→Cys) | 0.07 ± 0.02 | 70 |
| GST-XD4 (Ile338→Ser) | 0.4 ± 0.05 | 60 |
| GST-XD4 (Ile338→Ala) | 0.005 ± 0.002 | 50 |
| GST-XD4 (Ile338→Gly) | 0.005 ± 0.002 | 40 |
| GST-Fyn (wt) | 0.05 ± 0.02 | 100 |
| GST-Fyn (Thr339→Ala) | 0.005 ± 0.002 | 30 |
| GST-Abl (wt) | 0.3 ± 0.03 | 100 |
| GST-Abl (Thr334→Ala) | 0.03 ± 0.005 | 30 |
| Ser/Thr kinase | | |
| GST-CamKII (wt) | 60 ± 10 | 100 |
| GST-CamKII (Phe89→Ala) | 10 ± 2 | 90 |
| GST-CamKII (Phe89→Gly) | 0.5 ± 0.1 | 80 |
| Cdk2 (wt) | 50 ± 10 | 100 |
| Cdk2 (Phe80→Gly) | 0.16 ± 0.03 | 90 |
| P38 (wt) | 0.82 ± 0.2 | 100 |
| P38 (Thr106→Ala) | 0.027 ± 0.005 | 100 |
| P38 (Thr106→Gly) | 0.026 ± 0.005 | 100 |

than threonine at the corresponding position such as p38 γ and p38 δ . On the basis of our molecular modeling of Src family kinase Hck inhibition by PP1 and a sequence alignment, PP1 should inhibit wild-type p38 potently. The p38 kinase was expressed in *E. coli* [44] and the IC_{50} value for PP1 was determined. Indeed compared with wild-type CamKII α ($IC_{50} = 60 \mu\text{M}$) and Cdk2 ($IC_{50} = 50 \mu\text{M}$), PP1 inhibits wild-type p38 more potently ($IC_{50} = 0.82 \mu\text{M}$). Furthermore, when the Thr106 in p38 was mutated to alanine or glycine, both Thr106→Ala and Thr106→Gly have an IC_{50} of $\sim 0.025 \mu\text{M}$.

Although Ser/Thr kinases share less than 20% sequence identity with tyrosine kinases throughout the catalytic domains, they share high homology in the region directly surrounding the ATP-binding site, where the ATP-competitive inhibitors bind. This is most clearly demonstrated by the bis-indolemaleimide natural product staurosporine which is a potent ($IC_{50} \sim 10 \text{ nm}$) but nonselective protein kinase inhibitor [45]. Similarly, the flavone, quercetin is a general tyrosine kinase inhibitor ($IC_{50} \sim 24 \mu\text{M}$) that also inhibits protein kinase C ($IC_{50} \sim 31 \mu\text{M}$) [46]. For selective kinase inhibitors such as PP1, it is important to determine the residues that control binding to different kinases and

to determine whether the sequence determinants are highly context dependent. Mutagenesis of homologous residues in widely varying kinases might provide a way to determine these rules without the need to clone and express each member of the kinase superfamily. Clearly, more residues in addition to position corresponding to Ile338 in v-Src contribute to PP1 binding. These features could be uncovered by saturation mutagenesis of the active-site residues as well as determination of the crystal structure of PP1 bound to a Src family kinase.

Significance

Protein kinases play a central role in signal transduction cascades. The huge number, the tremendous redundancy and overlapping substrate specificity of protein kinases make the dissection of kinase signaling pathways a challenge to existing genetic and biochemical techniques. Highly selective small-molecule inhibitors of protein kinases are useful probes in deconvoluting the complex kinase signaling pathways, as well as being potentially therapeutic. Despite their value, highly specific kinase inhibitors have proven very difficult to obtain [47]. Here, we have studied a potent Src family selective kinase inhibitor, PP1, and identified a single residue that controls the ability of PP1 to potently inhibit protein kinases using molecular modeling and sequence analysis. The amino acid corresponding to Thr338 in c-Src is in close contact with the C3 phenyl ring of PP1. Kinases with residues larger than threonine at this position do not bind PP1 due to a steric clash with this residue. Mutation of the homologous residue to alanine or glycine in distantly related kinases such as CamKII α and CDK2 result in potent inhibition of these kinases by PP1. Understanding the structural basis of inhibitor specificity will help us to design more potent and more selective protein kinase inhibitors.

Materials and methods

Mutagenesis, protein expression and sequence comparison

Overlap extension PCR was used to make GST-XD4 mutants [48]. Pfu polymerase (Stratagene) was used according to the manufacturer's instructions. Eighteen synthetic oligonucleotides were used to generate GST-XD4 and GST-Fyn mutant kinases: primer 1 (5'-TTTGGATC-CATGGGGAGTAGCAAGAGCAAG), primer 2 (5'-TTTGAATTCCTA-CTCAGCGACCTCCAACAC), primer 3 (5'-CTACATCGTCTTTGAG-TACATGAG), primer 4 (5'-CTCATGTACTCAAAGACGATGTAG), primer 5 (5'-CTACATCGTCATGGAGTACATGAG), primer 6 (5'-CTC-ATGTACTCCATGACGATGTAG), primer 7 (5'-CTACATCGTCACTG-AGTACATGAG), primer 8 (5'-CTCATGTACTCAGTGACGATGTAG), primer 9 (5'-CTACATCGTCTGTTGAGTACATGAG), primer 10 (5'-CT-CATGTACTCAAAGACGATGTAG), primer 11 (5'-CTACATCGTCTG-CGAGTACATGAG), primer 12 (5'-CTCATGTACTCGCAGACGATG-TAG), primer 13 (5'-CTACATCGTCAGCGAGTACATGAG), primer 14 (5'-CTCATGTACTCGTCTGACGATGTAG), primer 15 (5'-CTACATCG-TCGCTGAGTACATGAG), primer 16 (5'-CTCATGTACTCAGCGA-CGATGTAG), primer 17 (5'-CTACATCGTCTGGGGAGTACATGAG), primer 18 (5'-CTCATGTACTCCCCGACGATGTAG). Primer 1 includes a BamHI cloning site and primer 2 has an EcoRI cloning site. Primer 3 and Primer 4 contain the Ile338→Phe mutation. Primers 5–18 contain the Ile338 to other mutations (see below). GST-XD4 (Ile338→Phe)

was made using primer 1, 2, 3 and 4 with GST-XD4 plasmid as the template. In the first PCR, primer 1 and primer 4 were used to produce one fragment and primer 2 and primer 4 were used to produce another fragment. In the second round of PCR, these two fragments were annealed and extended to form the full-length GST-XD4 (I338F) gene. The PCR products were digested with BamHI and EcoRI and ligated into similarly digested pGEX-KT. The resultant plasmid was transformed into the *E. coli* strain DH5 α . GST-XD4 (Ile338→Met, Ile338→Thr, Ile338→Val, Ile338→Cys, Ile338→Ser, Ile338→Ala and Ile338→Gly) were made in the same way except for use of corresponding internal primers (5 and 6 for Ile338→Met, 7 and 8 for Ile338→Thr, 9 and 10 for Ile338→Val, 11 and 12 for Ile338→Cys, 13 and 14 for Ile338→Ser, 15 and 16 for Ile338→Ala, 17 and 18 for Ile338→Gly). The GST-Abl (containing SH2 and SH1 domains) was from Y.J. Wang's lab (Department of Biology and Center for Molecular Genetics, UC at San Diego). GST-Abl (T334A) was constructed in the same way as GST-XD4 mutants except that the primers used were T-120A1: (5'-CTACATAATCGCTGAGTTCATG) and T-120A2: (5'-CA-TGAACTCAGCGATTATGTAG). Wild-type CamKII α was cloned into the pGEX-4T-1 vector from pBluscript CamKII α plasmid with the following primers: primer camk-1: (5'-CCC GGATCCATGGCTACCAT-CACCTGCAC). Primer camk-2: (5'-CCCGAATTCCTCAATGCGG-CAGGACGGAGG). Primer camk-1 contains a BamHI site and primer camk-2 contains an EcoRI site. CamKII α mutants were made using QuickChange site-directed mutagenesis kit (Stratagene) according to the manufacturer's instruction. The primers used to make CamKII α (Phe89→Ala) were: (5'-CCACTACCTTATCGGCGATCTGGTTACT-GG) and (5'-CCAGTAACCAGATCGCCGATAAGGTAGTGG). The primers used to make CamKII α (Phe89→Gly) were: (5'-CCACTACCT-TATCGGCGATCTGGTTACTGG) and (5'-CCAGTAACCAGATCGG-CGATAAGGTAGTGG). The Cdk2 ORF was cloned from plasmid pVLCDK2NHA into pFBH6 vector of Sf9 insect cell expression system and expressed and purified as described previously [49]. The primers used to make Phe80→Gly Cdk2 were: (5'-GATCTTGGTGCAGGAA-TTCAACCAACCAGGTAGAG) and (5'-CTCTACCTGGTTGGTGAATT-CCTGCACCAAGATC). The cDNA encoding 6-His p38 was generously provided by M.H. Cobb (University of Texas Southwestern Medical Center) in a pT7-5 construct. The Thr106→Ala and Thr106→Gly p38 mutants were generated by QuickChange site-directed mutagenesis kit (Stratagene) on the pT7-5 construct according to the manufacturer's instruction with the following primers: 5'-CCCCCATGAGATGGGGCCACCAGGTACAC-3', 5'-GTGTACTG-GTGGCCCATCTCATGGGGGC-3' (Thr106→Ala) and 5'-GCCCC-CATGAGATGGCCACCAGGTACACGTC-3', 5'-GACGTGTACTG-GTGGCCCATCTCATGGGGGC-3' (Thr106→Gly). All mutants were sequenced over the entire coding region (oligonucleotide sequencing and synthesis facility, Princeton University). Construction of GST-Fyn (wt) and GST-Fyn (Thr339→Ala) were described previously [50]. Expression and purification of v-Src, Abl and CamKII α wild-type and mutant kinases were carried out in DH5 α as described by Xu *et al.* [36] with the exception that the cells were stored at 4°C overnight before centrifugation and lysis by French press (overnight storage is essential for producing highly active kinases). Expression and purification of p38 wild-type and mutant kinases were carried out in BL21 (DE3) pLysS *E. coli* (Stratagene) as previously described by Khokhlatchev *et al.* [44].

In vitro kinase assays

General. *In vitro* kinase peptide assays were carried out in the presence of (2 μ Ci) 10 nM [γ -³²P] ATP (6000 Ci/mmol, NEN) in 30 μ l. (Typical kinase assays used to measure inhibitor IC₅₀ values utilize 10–15 μ M [γ -³²P] ATP [51]. Under such conditions it is difficult to compare IC₅₀ values for mutant kinases which have different K_M values for ATP {IC₅₀ = K_i (1 + [ATP]/K_{M,ATP})}. We used a low concentration of [γ -³²P] ATP (10 nM) such that IC₅₀ values determined would not be greatly influenced by the different K_M values for different protein kinases and mutants, typically 10–100 μ M for most wild-type kinases and 70–500 μ M for the mutant kinases described here [50]). Reaction mixtures (25 μ l) were spotted onto a phosphocellulose disk, immersed

in 10% HOAc, and washed with 0.5% H₃PO₄, once with 20 ml of acetone, followed by scintillation counting with a Beckman liquid scintillation counter. IC₅₀ values for PP1 were determined by measuring the counts per minute (cpm) of ³²P transferred to a phosphoacceptor substrate (given below). The transfer of ³²P was measured by standard scintillation counting. IC₅₀ is defined to be the concentration of inhibitor at which the cpm was 50% of the control (no inhibitor) disk. When the IC₅₀ fell between two measured concentrations, it was calculated based on the assumption of an inversely proportional relationship between inhibitor concentration and cpm between the two data points.

The percent activity (Table 2, last column) was determined by simultaneous purification of the wild-type and the mutant kinases. The percent activity of mutant kinase relative to the wild-type kinase was reported as a ratio of v₀ [cpm in the presence of mutant kinase]/ v₀ [cpm in the presence of wild type kinase], under the assay conditions for each kinase described above, in the absence of PP1.

XD4 and Fyn assays. Determination of IC₅₀ values for wild-type and mutant forms of XD4 was described previously [10]. Briefly, various concentrations of PP1 were incubated with 50 mM Tris (pH 8.0), 10 mM MgCl₂, 1.6 mM glutathione, 1 mg/ml of BSA, 0.1 mg/ml IYGE-FKKK (optimal Src/Fyn substrate), 3.3% DMSO and 0.1–50 nM GST-XD4 or GST-Fyn. It was confirmed that active GST-XD4 or GST-Fyn was present at very low concentrations (compared to PP1). We verified that PP1 IC₅₀ values remained constant over a 25-fold range in kinase concentration bracketing the standard assay conditions for the most potently inhibited mutants.

Abl assay. Various concentrations of PP1 were incubated with 50 mM Tris (pH 8.0), 10 mM MgCl₂, 1.6 mM glutathione, 1 mg/ml of BSA, 0.1 mg/ml EAIYAAPFAKKK (optimal Abl substrate), 3.3% DMSO, and 0.1–50 nM of GST-Abl. It was confirmed that active GST-Abl was present at very low concentrations (compared with PP1, see above).

Cdk2 assay. Various concentrations of PP1 were incubated with 50 mM Tris (pH 8.0), 10 mM MgCl₂, 1 mg/ml of BSA, 0.3 mg/ml Histone III-S, 3.3% DMSO, and 15 nM of CDK2/CyclinA.

CamKII α assay. The CamKII α assay kit was purchased from Upstate Biotechnology. Various concentrations of PP1 were incubated with 15 mM MOPS (pH 7.2), 18 mM β -glycerol phosphate, 0.75 mM sodium orthovanadate, 0.75 mM dithiothreitol, 0.75 mM CaCl₂, 20 mM MgCl₂, 0.5 μ M protein kinase A inhibitor peptide, 0.5 μ M protein kinase C inhibitor peptide, 125 μ M auto Camtide II (KKALRRQETVDAL), 10 μ g/ml Calmodulin, 0.05 μ M CamKII α for 20 min.

p38 assay. Various concentrations of PP1 were incubated with 50 mM Tris (pH 8.0), 10 mM MgCl₂, 1 mg/ml of BSA, 0.25 mM imidazole, 0.3 mg/ml myelin basic protein, 3.3% DMSO, and 0.1 nM of p38 MAPK.

Note added in proof

After this manuscript was submitted, two papers describing the crystal structures of PP1 bound to Hck [52] and Lck [53] have appeared.

Acknowledgements

We thank Melanie Cobb for the p38 expression system, Jean Wang for plasmid encoding c-Abl. The work was supported in part by the National Institute of Allergy/Infectious Disease (AI44009-1), the Searle Scholar's Program, the Pew Program in the Biomedical Sciences, the Cottrell Scholar's Program, and the Sloan Research Foundation. K.M.S. is a Glaxo Wellcome Scholar in Organic Chemistry.

References

- Isakov, N., Wange, R.L. & Samelson, L.E. (1994). The role of tyrosine kinases and phosphotyrosine-containing recognition motifs in regulation of the T cell-antigen receptor-mediated signal transduction pathway. *J. Leukoc. Biol.* **55**, 265-271.
- Perlmutter, R.M. (1995). Control of T cell development by non-receptor protein tyrosine kinases. *Cancer Surv.* **22**, 85-95.
- Martin, K.C., *et al.*, & Kandel, E.R. (1997). MAP kinase translocates into the nucleus of the presynaptic cell and is required for long-term facilitation in Aplysia. *Neuron* **18**, 899-912.
- Holm, N.R., Christophersen, P., Olesen, S.P. & Gammeltoft, S. (1997). Activation of calcium-dependent potassium channels in mouse [correction of rat] brain neurons by neurotrophin-3 and nerve growth factor. *Proc. Natl Acad. Sci. USA* **94**, 1002-1006.
- Hunter, T. (1994). 1001 protein kinases redux—towards 2000. *Semin. Cell Biol.* **5**, 367-376.
- Shokat, K.M. (1995). Tyrosine kinases: modular signaling enzymes with tunable specificities. *Chem. Biol.* **2**, 509-514.
- Thomas, S.M., Soriano, P. & Imamoto, A. (1995). Specific and redundant roles of Src and Fyn in organizing the cytoskeleton. *Nature* **376**, 267-271.
- Bolen, J.B., Rowley, R.B., Spana, C. & Tsygankov, A.Y. (1992). The Src family of tyrosine protein kinases in hemopoietic signal transduction. *FASEB J.* **6**, 3403-3409.
- Ilic, D., *et al.*, & Aizawa, S. (1995). Reduced cell motility and enhanced focal adhesion contact formation in cells from FAK-deficient mice. *Nature* **377**, 539-544.
- Bishop, A.C., Shah, K., Liu, Y., Witucki, L., Kung, C. & Shokat, K.M. (1998). Design of allele-specific inhibitors to probe protein kinase signaling. *Curr. Biol.* **8**, 257-266.
- Bishop, A.C., Kung, C., Shah, K., Witucki, L., Shokat, K.M. & Liu, Y. (1999). Generation of monospecific nanomolar tyrosine kinase inhibitors via a chemical genetic approach. *J. Am. Chem. Soc.* **121**, 627-631.
- Hunter, T. (1987). A thousand and one protein kinases. *Cell* **50**, 823-829.
- Norman, T.C., Gray, N.S., Koh, J.T. & Schultz, P.G. (1996). A structure-based approach to kinase inhibitors. *J. Am. Chem. Soc.* **118**, 7430-7431.
- Colas, P., Cohen, B., Jessen, T., Grishina, I., McCoy, J. & Brent, R. (1996). Genetic selection of peptide aptamers that recognize and inhibit cyclin-dependent kinase 2. *Nature* **380**, 548-550.
- Penninger, J.M., Wallace, V.A., Kishihara, K. & Mak, T.W. (1993). The role of p56lck and p59fyn tyrosine kinases and CD45 protein tyrosine phosphatase in T-cell development and clonal selection. *Immunol. Rev.* **135**, 183-214.
- Ezumi, Y., Shindoh, K., Tsuji, M. & Takayama, H. (1998). Physical and functional association of the Src family kinases Fyn and Lyn with the collagen receptor glycoprotein VI-Fc receptor gamma chain complex on human platelets. *J. Exp. Med.* **188**, 267-276.
- Fox, J.E. (1996). Platelet activation: new aspects. *Haemostasis* **26 Suppl 4**, 102-131.
- Kawakami, Y., Yao, L., Miura, T., Tsukada, S., Witte, O.N. & Kawakami, T. (1994). Tyrosine phosphorylation and activation of Bruton tyrosine kinase upon Fc epsilon RI cross-linking. *Mol. Cell Biol.* **14**, 5108-5113.
- Nishizumi, H., *et al.*, & Yamamoto, T. (1995). Impaired proliferation of peripheral B cells and indication of autoimmune disease in lyn-deficient mice. *Immunity* **3**, 549-560.
- Tsygankov, A.Y., Broker, B.M., Fargnoli, J., Ledbetter, J.A. & Bolen, J.B. (1992). Activation of tyrosine kinase p60fyn following T cell antigen receptor cross-linking. *J. Biol. Chem.* **267**, 18259-18262.
- Cooke, M.P., Abraham, K.M., Forbush, K.A. & Perlmutter, R.M. (1991). Regulation of T cell receptor signaling by a src family protein-tyrosine kinase (p59fyn). *Cell* **65**, 281-291.
- Levitski, A. (1996). SRC as a target for anti-cancer drugs. *Anticancer Drug Des.* **11**, 175-182.
- Showalter, H.D. & Kraker, A.J. (1997). Small molecule inhibitors of the platelet-derived growth factor receptor, the fibroblast growth factor receptor, and Src family tyrosine kinases. *Pharmacol. Ther.* **76**, 55-71.
- Hanke, J.H., *et al.*, & Connelly, P.A. (1996). Discovery of a novel, potent, and src family-selective tyrosine kinase inhibitor. *J. Biol. Chem.* **271**, 695-701.
- Bridson, S.J. & Watson, S.P. (1999). Evidence for the involvement of p59fyn and p53/56lyn in collagen receptor signalling in human platelets. *Biochem. J.* **338**, 203-209.
- Amoui, M., Draber, P. & Draberova, L. (1997). Src family-selective tyrosine kinase inhibitor, PP1, inhibits both Fc epsilonRI- and Thy-1-mediated activation of rat basophilic leukemia cells. *Eur. J. Immunol.* **27**, 1881-1886.
- Kapus, A., Szaszi, K., Sun, J., Rizoli, S. & Rotstein, O.D. (1999). Cell shrinkage regulates src kinases and induces tyrosine phosphorylation of cortactin, independent of the osmotic regulation of Na⁺/H⁺ exchangers. *J. Biol. Chem.* **274**, 8093-8102.

28. Howlett, C.J., Bisson, S.A., Resek, M.E., Tigley, A.W. & Robbins, S.M. (1999). The proto-oncogene p120(Cbl) is a downstream substrate of the Hck protein-Tyrosine kinase. *Biochem. Biophys. Res. Commun.* **257**, 129-138.
29. Bishop, A.C., Moore, D., Scanlan, T.S. & Shokat, K.M. (1997). Screening a hydroxystilbene library for selective inhibition of the B cell antigen receptor kinase cascade. *Tetrahedron* **53**, 11995-12004.
30. Siekierka, J.J. & Sigal, N.H. (1992). FK-506 and cyclosporin A: immunosuppressive mechanism of action and beyond. *Curr. Opin. Immunol.* **4**, 548-552.
31. Sigal, N.H. & Dumont, F.J. (1992). Cyclosporin A, FK-506, and rapamycin: pharmacologic probes of lymphocyte signal transduction. *Annu Rev. Immunol.* **10**, 519-560.
32. Zent, R., Ailenberg, M. & Silverman, M. (1998). Tyrosine kinase cell signaling pathways of rat mesangial cells in 3-dimensional cultures: response to fetal bovine serum and platelet-derived growth factor-BB. *Exp. Cell Res.* **240**, 134-143.
33. Frye, S.V. (1999). Structure-activity relationship homology (SARAH): a conceptual framework for drug discovery in the genomic era. *Chem. Biol.* **6**, R3-R7.
34. Siccheri, F., Moarefi, I. & Kuriyan, J. (1997). Crystal structure of the src family tyrosine kinase hck. *Nature* **385**, 602-609.
35. Garcia, P., Shoelson, S.E., George, S.T., Hinds, D.A., Goldberg, A.R. & Miller, W.T. (1993). Phosphorylation of synthetic peptides containing Tyr-Met-X-Met motifs by nonreceptor tyrosine kinases *in vitro*. *J. Biol. Chem.* **268**, 25146-25151.
36. Xu, B., Bird, G.V. & Miller, T.W. (1995). Substrate specificities of the insulin and insulin-like growth factor 1 receptor tyrosine kinase catalytic domains. *J. Biol. Chem.* **270**, 29825-29830.
37. Liu, Y., Shah, K., Yang, F., Witucki, L. & Shokat, K.M. (1998). Engineering Src family protein kinases with unnatural nucleotide specificity. *Chem. Biol.* **5**, 91-101.
38. Wang, J.Y. & Baltimore, D. (1985). Localization of tyrosine kinase-coding region in v-abl oncogene by the expression of v-abl-encoded proteins in bacteria. *J. Biol. Chem.* **260**, 64-71.
39. Colbran, R.J. (1992). Regulation and role of brain calcium/calmodulin-dependent protein kinase II. *Neurochem. Int.* **21**, 469-497.
40. De Bondt, H.L., Rosenblatt, J., Jancarik, J., Jones, H.D., Morgan, D.O. & Kim, S.H. (1993). Crystal structure of cyclin-dependent kinase 2. *Nature* **363**, 595-602.
41. Wilson, K.P., *et al.*, & Su, M.S. (1997). The structural basis for the specificity of pyridinylimidazole inhibitors of p38 MAP kinase. *Chem. Biol.* **4**, 423-431.
42. Evers, P.A., Craxton, M., Morrice, N., Cohen, P. & Goedert, M. (1998). Conversion of SB 203580-insensitive MAP kinase family members to drug-sensitive forms by a single amino acid substitution. *Chem. Biol.* **5**, 321-328.
43. Fox, T., *et al.*, & Wilson, K.P. (1998). A single amino acid substitution makes ERK2 susceptible to pyridinyl imidazole inhibitors of p38 MAP kinase. *Protein Sci.* **7**, 2249-2255.
44. Khokhlatchev, A., Xu, S., English, J., Wu, P., Schaefer, E. & Cobb, M.H. (1997). Reconstitution of mitogen-activated protein kinase phosphorylation cascades in bacteria. Efficient synthesis of active protein kinases. *J. Biol. Chem.* **272**, 11057-11062.
45. Ruegg, U.T. & Burgess, G.M. (1989). Staurosporine, K-252 and UCN-01: potent but nonspecific inhibitors of protein kinases. *Trends Pharmacol. Sci.* **10**, 218-220.
46. Kang, T.B. & Liang, N.C. (1997). Effect of quercetin on activities of protein kinase C and tyrosine protein kinase from HL-60 cells. *Chung Kuo Yao Li Hsueh Pao* **18**, 374-376.
47. Bishop, A.C. & Shokat, K.M. (1999). Acquisition of inhibitor-sensitive protein kinases through protein design. *Pharmacol. Ther.* **1**-10.
48. Reikofski, J. & Tao, B.Y. (1992). Polymerase chain reaction (PCR) techniques for site-directed mutagenesis. *Biotech. Adv.* **10**, 535-554.
49. Gray, N.S., *et al.*, & Schultz, P.G. (1998). Exploiting chemical libraries, structure, and genomics in the search for kinase inhibitors. *Science* **281**, 533-538.
50. Liu, Y., Shah, K., Yang, F., Witucki, L. & Shokat, K. (1998). Engineering src family protein kinases with unnatural nucleotide specificity. *Chem. Biol.* **5**, 91-101.
51. Meijer, L. (1995). Chemical inhibitors of cyclin-dependent kinases. *Prog. Cell Cycle Res.* **1**, 351-363.
52. Schindler T., Siccheri F., Pico A., Gazit A., Levitzki A., Kuriyan J. (1999). Crystal structure of Hck in complex with a Src family-selective tyrosine kinase inhibitor. *Mol. Cell* **3**, 639-648.
53. Zhu X, *et al.*, & Morgenstern KA. (1999). Structural analysis of the lymphocyte-specific kinase Lck in complex with non-selective and Src family selective kinase inhibitors. *Structure* **7**, 651-661.

Because **Chemistry & Biology** operates a 'Continuous Publication System' for Research Papers, this paper has been published via the internet before being printed. The paper can be accessed from <http://biomednet.com/cbiology/cmb> - for further information, see the explanation on the contents pages.